Excited State Dynamic-Node Diffusion Monte Carlo Simulations

Douglas R. Beck, Graeme Henkelman, and R. O. Watts*

School of Chemistry, University of Melbourne, Parkwille, Vic 3052

Department of Chemistry, University of Washington, Seattle, WA 98195

(June 12, 1997)

Abstract

We describe a novel diffusion Monte Carlo algorithm for calculating excited state quantum wave functions and energies which requires little prior knowledge of nodal surfaces. An initial, guessed, nodal surface is refined during the simulation by balancing populations in different nodal regions. The method requires constraints on an excited system to ensure that it does not relax to a less energetic mode. So long as such constraints can be found, the method will allow complex nodal surfaces to evolve dynamically, and accurate energies of excited systems to be calculated. Results are given for a simple vibrational system, ro-vibrational states of hydrogen fluoride, and excited vibrational states of the water molecule.

* Present address: BHP Melbourne Laboratories, 245 Wellington Road, Mulgrave, Vic 3170, Australia.

S.S. Xantheas (ed.), Recent Theoretical and Experimental Advances in Hydrogen Bonded Clusters, 155–185.

I. INTRODUCTION

The purpose of this paper is to investigate the use of the Diffusion Monte Carlo (DMC) method to simulate excited quantum systems. It builds upon work reviewed by Suhm and Watts [1]; this review also provides a good general discussion of quantum Monte Carlo method as applied to molecular vibrations Many of the underlying ideas were introduced by Anderson [2], and by Ceperley [3], to mention only a two contributors. Excited state simulations pose several challenges to the Monte Carlo method, especially near the nodal surface itself. Our goal is to simulate solutions to the time independent Schrödinger equation without having to predetermine the position of nodal surfaces, and with minimal foreknowledge of their complexity. The techniques apply especially to systems in which particles move in strongly coupled analytical potentials, or potentials which are perturbed from a simple form, thus making exact solutions difficult to calculate. In such cases, if it is possible to make an initial guess of the wave function, the methods discussed here provide an algorithm for relaxing the system an accurate solution.

II. DIFFUSION MONTE CARLO

The fundamental equation for DMC is obtained from the time-dependent Schrödinger equation by transforming to imaginary time, $\tau = it$, to give the diffusion equation with the addition of a first order rate term in the form of the potential (in atomic units where $\hbar = 1$)

$$\frac{\partial}{\partial\tau}\psi(\vec{r},\tau)=\left[\frac{1}{2}\sum_{\alpha}\left(\frac{1}{m_{\alpha}}\nabla_{\alpha}^{2}\right)-V(\vec{r})\right]\psi(\vec{r},\tau). \qquad (1)$$

The diffusion coefficient, D_{α}, for particle α is 1/(2).

Following standard procedures, the formal solution to this equation can be written as an expansion in terms of the eigenvalues, E_n and eigenfunctions, φ_n, of the Hamiltonian

$$\psi(\vec{r},\tau)=\sum_n C_n e^{-E_n\tau}\varphi_n(\vec{r}). \tag{2}$$

The eigenfunction components of the initial wave function either grow or decay exponentially along the imaginary time axis, depending upon the sign of the corresponding eigenvalue. Assuming that the initial state contains the ground state as one of its components, $\phi_n(\vec{r})$ an be generated by propagating the system to large values of imaginary time.

The DMC algorithm implements the propagator in its configuration space representation using a split operator approximation described in several papers [1-3].

$$e^{-\hat{H}\tau}=e^{-\hat{T}\tau}e^{-\hat{V}\tau} \tag{3}$$

Since this form ignores the non-commuting nature of $\hat{T}$ and $\hat{V}$, it is only valid to first order in τ. Propagation along the imaginary time axis proceeds iteratively by evolving the system through small increments in time, $\Delta\tau$. The kinetic propagator is represented as random walks in configuration space by a large number of replica systems. In the time interval $\Delta\tau$, each replica is displaced by a Gaussian distributed random step with standard deviation $\sigma=\sqrt{2D_n\Delta\tau}$. The potential operator is represented by a multiplicative factor $\exp\left((-VC\vec{r})\Delta\tau\right)$ in the weight of each replica [1].

III. MEASURING THE ENERGY OF A SYSTEM

In the standard DMC algorithm introduced by Anderson [2], a reference energy is introduced and adjusted in order to stabilize the population of replicas. The potential is measured from this reference energy and at each time step the replica weights are multiplied by $\exp\left(-\left(V-E_R\right)\Delta\tau\right)$ instead of $\exp\leq(-V\Delta\tau)$. The energy E_n in Eq. 2 is replaced by E_n-E_R, so that if $E_R = E_n$ for some n, the corresponding eigenfunction will be stabilized, and all higher energy components will decay exponentially in increasing τ. Assuming no other constraints are imposed on the simulation, the reference energy which stabilizes the

population is the ground state energy. More generally, it is the lowest eigenvalue consistent with the applied constraints. It follows that once equilibrium is reached, the lowest eigenenergy can be estimated by averaging the reference energy over a long sequence of time steps.

Another way to look at the effect of the reference energy, is that when the population is stabilized, the equilibrium wave function $\psi(\bar{r})$ is being rescaled at each time step by the factor $\exp(E_R \Delta\tau)$ A better method of finding E_R, then, is to pick a constant value for the total population weight, $\int \psi(\bar{r}) d\bar{r}$, and record the multiplication factor R required to renormalize the population weight at each time step. The reference energy is then given by

$$E_R = -\ln R / \Delta\tau \tag{4}$$

and the eigenenergy can be estimated by averaging E_R over long τ.

Our approach to simulating quantum excited states is to divide the system into separate lobes or regions of configuration space. It is important to be able to independently estimate the eigenenergy of these regions. The reference energy is a valid eigenenergy estimator even if computed in a partial region of a stationary wave function, since the rescale factor R must be constant throughout. If R were different in one region than another within $\psi(\bar{r})$, the wave function would be changing shape and hence not be stationary. The concept of a local eigenenergy, E_{loc}, can be defined as the equilibrium reference energy in a local region of $\psi(\bar{r})$.

The DMC method yields wave functions for a stationary state. In this case the average potential energy for the population of replicas - which is not the same as the quantum expectation value - is also an estimator of the corresponding eigenenergy. This result can be shown by integrating the Schrödinger equation over some region of configuration space:

$$\int \hat{H}\psi(\vec{r})d\vec{r} = E\int \psi(\vec{r})d\vec{r} = \sum_{\alpha}\frac{1}{2m_{\alpha}}\int \Delta_{\alpha}^{2}\psi(\vec{r})d\vec{r} - \int V(\vec{r})\psi(\vec{r})d\vec{r} \qquad (5)$$

Using Green's theorem, the integration of $\nabla_{\alpha}^{2}\psi(\vec{r})$ *is* obtained by integrating the derivative of $\psi(\vec{r})$ across the nodal surface:

$$E = \frac{1}{\int \psi(\vec{r})d\vec{r}}\sum_{\alpha}\frac{1}{2m_{\alpha}}\int_{\tau}\nabla_{\alpha}\psi(\vec{r}).d\vec{s} - \overline{V} \qquad (6)$$

The term

$$\overline{V} = \frac{\int V(\vec{r})\psi(\vec{r})d\vec{r}}{\int \psi(\vec{r})d\vec{r}} \qquad (7)$$

is the average potential energy of the wave function within the region. In the case that the region of configuration space containing $\psi(\vec{r})$ is unbounded, the wave function and its derivative must vanish as $\vec{r} \to \infty$, and the average potential energy of the entire wave function is the eigen energy of the stationary state.

Practically, $\overline{V}$ is often a poor way to estimate the energy of the system. In the case of an antisymmetric wave function for example, the expression for $\overline{V}$ may be a limiting ratio of two vanishing functions. Such quantities are very difficult to estimate numerically. In eigenstates with the same number of positive and negative lobes, $\int \psi(\vec{r})d\vec{r}$ certainly becomes small, increasing any numerical uncertainty in a statistical estimate of $\overline{V}$. Furthermore, in our algorithm, to be discussed below, it is necessary to determine independently the local energy in each lobe of the wave function. In order to use $\overline{V}$ as a measure of local energy one must be able to estimate the gradient of the wave function at the boundary of the region of interest, as shown in Equation 6. This derivative is a difficult quantity to accurately measure from a

Monte Carlo simulation. For these reasons, $\bar{V}$ is not used as an energy estimator in the following techniques.

IV. DYNAMIC NODE SIMULATION

All methods for obtaining excited state wave functions and energies rely on applying constraints such that the ground state of the constrained system corresponds to an excited state of the unmodified system. Several constraint methods are available, but each has its limitations and much research is still needed to develop more efficient and flexible algorithms.

Some current techniques include orthogonalization constraints [4,5], basis-set approaches coupled with quantum Monte Carlo for integral sampling [3], fixed node simulations [2], and nodal relaxation schemes [3]. In the present paper we develop an effective dynamic nodal constraint approach.

Nodal constraint DMC algorithms rely on a prior knowledge of the location of the node, and the difficulty of determining nodal surfaces is its principal limitation. An absorbing wall which confines the replicas to one region of configuration space is placed on the nodal surface. During the simulation, replicas diffuse, grow, and decay as usual but with the modification that those which traverse the node meet their demise; their weight is set to zero. When such crossings occur, the replica with the largest weight has its weight reduced by half and is copied into the memory location of the deleted replica. The two replicas then propagate independently. Since the DMC algorithm is a finite time step simulation, there is a non-zero probability that a replica will cross and then recross a node within a single time-step. Anderson [2] shows that the recross probability is:

$$P_{\text{recross}} = \exp\left(\frac{-4|\vec{r}||\vec{r}'|}{2\sigma^2}\right), \tag{8}$$

where $\sigma^2 = 2D_t\Delta\tau$ is the variance and D_t the diffusion coefficient along an axis normal to the node, and $|\vec{r}|$ and $|\vec{r}'|$ are the projected distances of the replicas onto this axis before and after the time step. At the end of every time step, the weight of each replica is attenuated according to Equation 8. Without this correction, the recross error can be quite significant for DMC calculations.

The starting point for dynamic-node DMC is the simple fixed-node algorithm. Hard absorbing walls are placed along the guessed location of the nodal surface. However, rather than having one ensemble of replicas, multiple ensembles of replicas exist such that each ensemble is located in a separate region of configuration space, each region being bounded by a nodal surface. That is, every lobe in the wavefunction is modelled by its own ensemble of replicas. For example, two ensembles are established in a simulation of the first excited state of the one dimensional harmonic oscillator, one to the left and the other to the right of the node. The two ensembles are then treated as weakly interacting simulations. The key to the method is to monitor the difference between values of E_{loc} from the two lobal simulations as a function of the location of the node. If the location is correct, the estimators will give identical results and no adjustment is necessary. However, if the location is incorrect, then values of E_{loc} predicted by the two ensembles will be different. The algorithm introduced here incorporates a scheme for changing the location of the node so as to converge the energy estimates from the two simulations.

Given the stochastic nature of DMC, the instantaneous values of E_{loc} will differ from time step to time step dueto fluctuations. It is good practice to average E_{loc} in each lobe of the wave function over several time steps, before shifting the position of the node, in order to reduce statistical fluctuations in its position. If the absorbing node is shifted too often, the simulation will over-estimate E_{loc} in all nodes by effectively reducing the volume which $\psi(\vec{r})$ can occupy.

To determine how to move the node, we used a particle in a one-dimensional box as the starting point. The ground state energy for such a system is:

$$E = \frac{\pi^2 \hbar^2}{2mL^2} \tag{9}$$

where L is the length of the box and m is the mass of the particle. If the edge of the box is moved a small distance Δx, the expression for the change in energy is

$$\Delta E = -\frac{\pi^2 \hbar^2}{mL^3} \Delta x \tag{10}$$

Hence for small distances, the change in energy is directly proportional to the nodal displacement.

Suppose we now insert an absorbing barrier at a point X in the box, to give two boxes of length X and L-X. Simulations are carried out in each box to give local eigenvalues E_{loc}^1 and E_{loc}^2, say. We now move the location of the absorbing barrier by an amount

$$\Delta x = -\varepsilon\left(E_{loc}^1 - E_{loc}^2\right) = -\varepsilon \Delta E \tag{11}$$

where the proportionality constant, ε, *is* determined in trial DMC simulations to optimize convergence of the energies. Simulations of E_{loc} in the new boxes will give values which are closer. Proceeding in this way, we will reach the condition $E_{loc}^1 = E_{loc}^2$, at which point the boxes will be of equal length, $L/2$. This value of E_{loc} corresponds to the first excited state energy of the original box.

The elements for a dynamic-node algorithm are now in place. With the nodal surface placed in some initial configuration, the replicas diffuse, and grow or decay, as in a simple DMC simulation. At the end of a short simulation, the nodal displacement is calculated according to the above rule. The difference in energy estimators between ensembles on

opposite sides of each nodal surface is determined, multiplied by ε, and the surface is then moved along the surface normal by an amount given by this product. The negative sign in the algorithm indicates that the node moves in a direction such that the higher energy lobe expands and the lower energy lobe contracts. This algorithm makes three assumptions about the nodal characteristics: (1) the nodal surface has a simple shape, such as a line, plane, or sphere; (2) that the node moves such that this shape is conserved; and (3) that the direction in which to move the node is such that one volume always contracts and the other expands. A simple system in to these conditions apply is a one dimensional Morse oscillator. We consider more complex surfaces, where the shape of the surface must also be allowed to change, later in the paper.

V. ONE DIMENSIONAL MORSE OSCILLATOR

The Morse oscillator potential

$$M(x) = D\left(e^{-\alpha x} - 1\right)^2 \tag{12}$$

was chosen because it is not symmetric and is often a good approximation to the radial potential between interacting particles. Each simulation used the same parameters. In atomic units, the well depth $D = 10.0$, the exponential factor $\alpha = 0.200$, and the particle mass, $m = 1$. The eigenvalues and eigenfunctions of the Morse potential are known analytically, and the nodal positions and energies of the ground state and first three excited states are given in Table I for comparison with the DMC results.

To initiate the excited state simulations, nodal points were chosen at intervals of 1 atomic unit spaced around $x = 0$. Replicas were then generated in a Gaussian distribution about the origin. If a replica was found to be in a position $x < N_1$, the coordinate of the first node, it was considered to lie in Region I. If it was in a region $N_1 < x < N_2$, where N_2 is the coordinate of the second node, it was labelled as Region II, and so on. The process of positioning replicas

was continued until there were 2,000 replicas in each lobe of the wave function being simulated. All replicas were given the same initial weight, 0.0005, so that every lobe had the same integrated weight of 1.0. Neither the initial positions of the nodes nor those of the replicas were of any great consequence as all change dramatically over the first few iterations, bringing the nodes and the lobes of the wave function close to the analytical solution. It is much more difficult to accurately estimate the eigenenergy.

A time step of 0.02 E_b^{-1} was chosen for all Morse simulations. Finding a good value for the quantity $\Delta\tau$ is an art. If it is too small, the simulation takes a long time to reach equilibrium. Even when the system is in steady state, obtaining accurate values of the energy requires the averaging of many independent configurations of the replicas. For overly small time steps, it takes many iterations to achieve a new independent replica arrangement and little is learned from the many calculations between these configurations. On the other hand, if the time step is too large, the split operator approximation described in Eq. 3 starts to become invalid and the simulation will yield an artificially inflated estimate of the energy. A good rule of thumb is that for a valid simulation, one should not see a variation in the energy estimation as $\Delta\tau$ is reduced.

The simulations were done in three stages. A first run of 5,000 iterations, in which ε was chosen to be 2 x 10^{-3}, served to bring the wave functions and their nodes close to their final configurations. The recross correction, Eq. 8, was applied for higher accuracy. Local energies, E_{loc}, were estimated for each lobe independently, by calculating the constant required to renormalize the sum of the weights of replicas in that lobe, Eq. 4. To reduce coupling between random fluctuations in the energy estimators and movement of the nodes, the energy, E_{loc}, in each lobe was averaged over 5 time steps before moving the nodes. Every node was then shifted according to the difference in the average values of E_{loc} for the lobes on either side of the node. For example, for the first node $\Delta N_1 = -\varepsilon \left(E_{loc}^{I} - E_{loc}^{II}\right)$ according to Equation 11.

In the second stage, simulations were run for each lobe to obtain a time averaged value of the energy E_{loc} over 20,000 iterations. To start the simulations, each replica was split into five, yielding 10,000 replicas in each lobe. The energies averaged over 10 time steps before shifting the nodes, and ε was increased to 5×10^{-3} to compensate for much smaller differences in the energies between lobes. This stage of the simulation resulted in well-converged energies for the ground and first excited states, but the second and third excited state energies had not yet reached equilibrium. A third stage of simulation was then done on each system using the same parameters. The results of the final simulations are displayed in Table 1 for comparison with the exact values. Uncertainties were computed by finding the number of independent configurations within the 20,000 iterations and evaluating the standard error of the energies and nodal positions within this set.

Generally, the DMC results agree with the exact values to within their statistical uncertainties. In the third excited state, there is some evidence of the DMC simulation overestimating the eigenenergy. This defect is due to random movements in the nodal positions, which effectively reduces the volume of the lobes. These fluctuations also bias the position of the nodes, because a change in the volume of a lobe has more effect in regions where the potential is changing rapidly. The simulation for the third excited state was repeated using a smaller value of ε and a larger number of timesteps between nodal shifts. These changes removed the slight discrepancy between the DMC and the theoretical energies.

As can be seen from Table 1 the simulations yielded good values of the Morse potential eigenenergies and nodal positions. However, a histogram of the final wave function constructed from the global distributions would not be correct. The normalization constraint applied at the end of every timestep ensured that the sum of population weights in every lobe was the same. If wavefunctions are needed, this is incorrect on two counts. First, it is the sum of the squares of a population histogram that must sum to unity, and second the sum of

weights within different lobes in general will be different. In order to generate the wave function, the populations in each lobe must be scaled such that the slopes across each node are continuous, and that $\int \psi(\vec{r})^2 d\vec{r} = 1$. This process is described in more detail below. With these procedures applied, Figure 1 shows the ground state and the first two excited states of the Morse Oscillator.

VI. HYDROGEN FLUORIDE

As a second test of the dynamic node algorithm, the diatomic molecule *HF* was modelled in three dimensions. Simulations were conducted in the center-of-mass frame. The Hamiltonian operator, in radial coordinates, has the following form:

$$-\frac{1}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{r^2}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\theta\partial\phi^2}\right) + V(r) \qquad (13)$$

where $\mu = 1744.68$ for *HF*. The potential used was a Morse oscillator:

$$V(r) = V_e\left(1 - e^{-a(r-r_e)}\right)^2 \qquad (14)$$

with $V_e = 0.18379$, $a = 1.29938$, and $r_e = 1.73278$ [5]. This choice of interaction allows us to make a ready comparison with exact results.

In order to compare DMC nodal positions and energies with analytical values, the rigid rotor approximation was used. The eigenfunctions of the Hamiltonian are of the form

$$\phi_{v,l,m}(r,\theta,\phi) = R(r)Y_l^m(\theta,\phi) \qquad (15)$$

with

$$R(r) = \frac{1}{r}X_v(r) \qquad (16)$$

where $X_v(r)$ are the eigenfunctions of the one dimensional Morse oscillator and $Y_l^m(\theta,\varphi)$ are the spherical harmonics. Within the rigid rotor approximation, the eigenvalues are the sum of the Morse vibrational energies and the rotational energies:

$$E_{v,l} = E_v + B2\pi l(l+1) \tag{17}$$

where B, the rotational constant for HF, is $1.5191 \times 10^{-5}\ E_h$.

In the following subsections we describe the simulations and summarize the results. The nodal positions are well determined within the rigid rotor approximation. Simulations were run for the following ro-vibrational states: (1) $v = 0$, $l = \ell$ 1, (2) $v = 0$, $l = 2$, (3) $v = 1$, $l = 0$ and (4) $v = 1$, $l = 1$.

For the first excited rotational state ($v = 0$, $l = 1$) there are three degenerate eigenvalues, each having one planar node, lying in the xy, xz, and yz planes respectively. The dynamic node simulation in this example had a nodal barrier that was constrained to be parallel to the xy plane. The nodal adjustment moved the entire plane along the $\hat{z}$-axis while keeping the plane parallel to the xy plane. Thus, the simulation was analogous to the one dimensional case with a point node.

The other excited rotational state to be simulated ($v = 0$, $l = 2$) has five degenerate eigenvalues. We simulated the state corresponding to the $2P_{xy}$ state where both nodes are planar and lie in the xz and yz planes. This creates four nodal regions, although only three are required to move the nodes; therefore, the three simulations were run in the quadrants ($-x,-y$), ($-x,y$), and (z,y), labelled I, II, and III respectively in this paper. As with the one-dimensional example, simulations of lobes I and II were coupled as were lobes II and III. The nodes were constrained to remain parallel to the axial planes and were thus moved along the $\hat{x}$ and $\hat{y}$ axes.

A simulation of the lowest energy vibrational state (v = 1, l = 0) requires a radial node, which is simply a sphere with its center at the center-of-mass. Movement of this node involved a simple expansion or contraction of the nodal radius by an amount Δr, while the center remained fixed. This breathing motion caused a volume change in the lobes that scaled as Δr^3, and raised the question as to whether or not Eq. 11 can be applied unchanged. We applied it nonetheless, and found that this algorithm was able to converge the energies, as shown in the following results. Most likely, there are more efficient ways to achieve convergence.

The lowest energy state with both vibrational and rotational excitation (v = 1, l = 1) has both a spherical node centered at the center-of-mass and a node parallel to the xy plane. As with the l = 2 example, three simultaneous simulations were run with the replicas placed in the following regions: (I) $-z$ hemisphere and inside the region bounded by the spherical node, (II) in the $+z$ hemisphere and inside the spherical node, and (III) in the $+z$ hemisphere and outside the spherical node. Simulations II and III were coupled and controlled the expansion and contraction of the radial node. Simulations I and II were coupled and controlled movement of the planar node: it moved as in previous examples, so that it was constrained to remain parallel with the xy plane and moved along the $\hat{z}$-axis.

All the results for HF are given in Table 2. The simulations were based on ensembles containing 10,000 replicas within each nodal region, a time step of $\Delta\tau = 0.5E_k^{-1}$, and local energies, E_{loc}, which were averaged over 20,000 time steps. The recross correction was applied to allow for the presence of the nodes and to achieve higher accuracy. In all cases, the nodes were given arbitrary initial positions; all converged to the correct location. Again, improved accuracy was obtained by reducing the nodal adjustment parameter and by decreasing the time step. In this simple system, dynamic-node DMC easily dealt with multiple nodes and nodes of different shapes.

VII. NODAL SURFACES OF UNKNOWN TOPOGRAPHY.

The results of the previous sections are of limited use for more complex nodal surfaces. It is not often that a multi dimensional problem has a nodal surface that can be parameterized with a single variable. In order to deal with more complicated nodal surfaces, additional constraints can be applied. The QMC algorithm used in the previous sections constrained E_{loc} be the same in every lobe. Nodal boundaries were adjusted dynamically to meet this constraint. Another property which can be employed as a constraint is that the slope of the wave function must be continuous across the nodal surfaces. By making use of this additional constraint, we are able to adjust the shape of the nodal surface as the simulation progresses.

We construct a trial surface which is composed of a set of small hyperplanes, s_k, which form a continuous surface. The derivatives taken along the surface are not necessarily continuous. During the simulation, these small hyperplanes are moved with respect to some origin according to two constraints. First, the full surface is displaced without change in shape in order to bring values of E_{loc} in all lobes to the same value. The basic algorithm for this step has been described. Secondly, the small hyperplanes are displaced along a direction normal to their surfaces in order to achieve continuous slopes across the whole nodal boundary. After making such an adjustment, the directions of the surface normals may need changing to ensure continuity along the nodal boundary.

In order to achieve continuity of slope, we note that in a properly normalised wavefunction, the flux of replicas crossing the nodal surface from the positive side is equal to that from the negative. In the present simulation, the constraint on E_{loc} *is* achieved by demanding that the population of replicas in all lobes is identical. We noted earlier that with this constraint, a simple histogram of replica positions does not give a continuous, normalized, wavefunction. This problem needs to be taken into account when trying to equalise fluxes. Note that the replica weights within a single node are correctly scaled with respect to each other.

Accordingly, we first calculate the fractional fluxes of particles crossing from positive to negative, f_k^+, and from negative to positive, f_k^-

$$f_k^+ = W_k^+ / \sum W_k^+ \tag{18}$$

$$f_k^- = W_k^- / \sum W_k^- \tag{19}$$

where $W_k^{+,-}$ are the weights of all replicas crossing surface element s_k in the time interval. We then compute the fractional fluxes

$$F_k^+ = \frac{f_k^+}{f_k^+ + f_k^-} \qquad F_k^- = \frac{f_k^-}{f_k^+ + f_k^-} \tag{20}$$

The element of surface s_k is then moved an amount Δr_k along its surface normal according to the formula

$$\Delta r_k = -\left(\varepsilon_E \Delta E_R + \varepsilon_F \left(F_k^- - F_k^+\right)\right) \tag{21}$$

The scaling factor ε_F is chosen to optimise convergence. Special constraints are needed at the extremities of the surface, and the second term is ignored for cases where no replicas cross from either direction. We have also explored modifications in which the planar surface s_k is replaced by a quadratic spline, allowing continuity along the nodal surface to be imposed more easily.

The new dynamic node algorithm solves one part of the excited state problem for DMC simulations, that of making the nodal surface flexible so it can relax into a fairly general shape. It does not solve the problem of simulating excited states of the same symmetry which have the same number of lobes as a state of lower energy. In such cases, additional orthogonality constraints must be imposed. To explore the algorithm, and its successes and limitations, we simulated excited vibrational states of the water molecule.

VIII. A THREE DIMENSIONAL SYSTEM: THE WATER MONOMER

The water molecule was simulated in a potential surface developed by Reimers and Watts [6] and later modified by Coker, Miller and Watts [7]. In this model, the potential is written in terms of three local mode coordinates:

$$s_1 = R_1 \cos\left[\frac{1}{2}(\theta - \theta_0)\right] - R_0 , \tag{22}$$

$$s_2 = R_2 \cos\left[\frac{1}{2}(\theta - \theta_0)\right] - R_0 , \tag{23}$$

$$s_3 = \left[\frac{(R_1 + R_2)}{R_0}\right] \sin\left[\frac{1}{2}(\theta - \theta_0)\right], \tag{24}$$

where R_1, R_2, and θ are the bond lengths and bond angle respectively; Ro and θ_0 are the corresponding equilibrium values. The potential surface is expressed as a sum of three Morse potentials and a simple quadratic coupling between s_1 and s_2:

$$V(s_1, s_2, s_3) = \sum_{i=1}^{3} D_i \left(1 - e^{\alpha_i s_i}\right)^2 + f_{12} s_1 s_2 \tag{25}$$

Parameters in the potential were determined by fitting a large basis set vibrational calculation to 56 vibrational levels for H_2O, D_2O, and HDO, and have the values $D_1 = D_2 = 0.20916$, $D_3 = 0.15660$, $\alpha_1 = \alpha_2 = 1.1331$, $\alpha_3 = 0.70600$, $R_0 = 1.8088$, $\theta_0 = 104.52^\circ$, and $f_{12} = -0.0067622$, all in atomic units.

For comparative purposes, approximate eigenstates for this potential surface can be constructed by assuming a) that the quadratic coupling term between s_1 and s_2 is small; and b) that the internal coordinates are rectilinear and the kinetic operator is diagonal. In this approximation, the molecule has three independent modes of vibration, and the full wave function can be expressed as a product of one-dimensional Morse oscillator states in s_1, s_2 and s_3. Reimers and Watts [6] have shown that the eigenfunctions obtained in this way are close to

those of the converged states. The rectilinear components of the local coordinates s_1 and s_2 are projections of the two hydrogen atoms onto the equilibrium bond vectors. The third coordinate, s_3, describes the relative displacement of the hydrogen atoms tangential to the bond vectors and corresponds to the bending mode in normal coordinates. Since we will describe the molecule in local coordinates, we note that symmetry constrains the wave functions to be symmetric or antisymmetric with respect to exchange of the coordinates s_1 and s_2.

Following Reimers and Watts [6], the wavefunctions will be described in terms of the quantum excitations in each of the local coordinates, $|n_1,n_2,n_3\rangle$. For example, if the molecule has one quantum of energy in the local mode s_3, the state is labelled $|0,0,1\rangle$. Under the assumptions described in the previous paragraph, a state $|n\rangle$ is the n^{th} eigenfunction of a Morse potential in the associated local coordinate. In the case of stretch excitations, s_1 and s_2, it is impossible to excite an identified O-H bond as the symmetry condition would be violated. For example, the state $|1,0,0\rangle$ is not a valid quantum state of the system whereas the (un-normalized) symmetry adapted states $|1,0,0\rangle_s = |1,0,0\rangle + |0,1,0\rangle$ and $|1,0,0\rangle_a = |1,0,0\rangle - |0,1,0\rangle$ are true eigenstates.

Figure 2 shows the low energy wavefunctions of the water monomer calculated using the basis states of Reimers and Watts [6]. The plots show projections onto the s_1s_2 plane for all states with less than three quanta in the stretch coordinates. Bend excitations in the s_3 coordinate do not alter the shape of these plots since s_3 is not coupled to the stretch coordinates by the symmetry constraint. The 3D wave function is the product of the wave functions shown in Figure 2 and the appropriate Morse oscillator function in the s_3 coordinate. It is interesting to note that the projections onto the s_1s_2 plane can show rather complicated nodal surfaces due to the symmetrization of the Morse states. It is this property of the wave

function that makes the water molecule a good test for the dynamic node Monte Carlo method.

Exact solutions to the full potential described in Equation 2 were calculated by Coker, Miller, and Watts [7] using a variational technique, starting from the reference states shown in Figure 2. The eigenenergies (< 9000 cm^{-1}) of the converged solutions are given in Table 3. Although the exact wavefunctions cannot be expressed precisely as a product of independent Morse states, they are described using the same quantum numbers. Differences between the reference and converged states arise from the coupling between s_1 and s_2 as well as off-diagonal kinetic energy terms, and are small.

Figure 2 can be used to begin a classification of wavefunctions in terms of geometry rather than symmetry. Both classifications are used in the following discussion. We see from the figures that the nodal surfaces are of two basic types, open and closed. A good example of a state with a single, open, nodal surface is $|1,0,0\rangle_a$, whose nodal surface is defined by the plane $s_1 + s_2 = 0$. Other states having a single, open nodal surface are $|0,0,1\rangle$ and $|1,0,0\rangle_s$. It is also possible for a state to have a single, closed, nodal surface. For example, the state $|2,0,0\rangle_s$ has a radial nodal surface with three maxima on the outer surface, as is seen in the figure. A common characteristic of the states with a single nodal surface is that there is one positive lobe and one negative lobe. A more complex geometric surface consists of the intersection of two open surfaces, for example the states $|2,0,0\rangle$ and $|1,1,0\rangle$. Geometric constraints of this type can be used to reduce the flexibility of the assumed nodal surfaces, although it is possible for open surfaces to become closed, and *vice versa* during the simulation. In addition to the geometric classification, the evolving nodal surfaces should also be subjected to appropriate symmetry constraints.

The DMC simulations were always conducted in stages. To begin, a fairly small number of replicas systems were simulated with a relatively large time step, to bring the simulation close to equilibrium. The conditions used during this process are relatively unimportant and will not be given for each simulation. For accurate calculations, between 1,000 and 5,000 replicas were simulated over 10,000 iterations with a time step of $2\,E_h^{-1}$. In the diffusion step, the atoms were displaced in laboratory-fixed cartesian coordinates and at the end of every time-step the axes were shifted to maintain the oxygen atom at the origin. After the simulation reached equilibrium 30,000 replicas were evolved over 20,000 iterations with a time step of $0.5\ E_h^{-1}$. We first simulated the ground state, $|0,0,0\rangle$,obtaining an average reference energy of $0.0210661\pm6\times10^{-6}\ E_h$ or $4623.5\pm1.4\ cm^{-1}$. The converged variational ground state energy reported by Coker *et al* [7] is $4623.02\ cm^{-1}$.

In order to simulation the excited states, the following symmetry coordinates were introduced to define nodal positions:

$$u = s_1 + s_2 \tag{26}$$

$$v = s_1 - s_2 \tag{27}$$

The antisymmetric stretch, $|1,0,0\rangle_a$, is the simplest excited state to simulate because its nodal surface is forced by symmetry considerations to lie in the plane $s_1 = s_2$. For any antisymmetric stateof the water molecule, the wave function by definition must change sign at $s_1 = s_2$ and hence its nodal surface must include this plane.

At first the $|1,0,0\rangle_a$ simulation was run without the recross probability correction. After very long runs and using a small time step, the energy converged to a value more than $200\ cm^{-1}$ lower than the converged variational value of $3755\ cm^{-1}$. This result led us to use the recross probability for all future simulations.

In order to compute the recross probability for a replica, it is necessary to know, on average, how far it will diffuse towards the nodal surface in a single time step. For the $|1,0,0\rangle_a$ state, the direction normal to the node is parallel to the vector $\vec{u}$. The diffusive displacements are calculated in cartesian space, and it is a tedious task to transform the kinetic operator into the local coordinates. To avoid this process, the diffusion length in the directions $\vec{u}$, $\vec{v}$, and $\vec{s}_3$ were determined by propagating the system and recording the value of the displacements in each of these directions. The diffusion length was calculated based upon the standard deviation of these displacements and found to be (in atomic units) $D_u = D_v = 5.6\times10^{-4}$ and $D_{s3} = 1.9\times10^{-4}$. It was also found that the measured diffusion lengths did not vary significantly with respect to position in the wave function, and they were therefore treated as constants. Using this correction, the $|1,0,0\rangle_a$ simulation was run with 10,000 replicas over 200,000 iterations of 0.1 E_h time step. An average excitation energy of 3751.8 ± 2.7 cm^{-1} was calculated in very good agreement with the variational value.

The $|1,0,0\rangle_a$ state belongs to the class of functions containing two lobes divided by an open nodal surface. The state for which the nodal surface is defined by the plane $u = 0$ is not the lowest energy configuration for systems within this class. The $|0,0,1\rangle$ and $|1,0,0\rangle_s$ states also have single, open, nodal surfaces and have lower eigenenergies. However, symmetry constraints allow the $|1,0,0\rangle_a$ state, to be simulated directly using the constraint $u = 0$. When the planar nodal surface v = 0 was allowed to translate and rotate in order to balance weight fluxes, but with no other constraints, the simulation converged towards the lowest energy in its geometric class which is the $|0,0,1\rangle$ bend mode.

The first symmetric bend state, $|0,0,1\rangle$ is the lowest excited vibrational state of the molecule. Using the approximations described above, a nodal plane was placed initially at $s_3 = 0$. Ensemble I was established with values of $s_3 < 0$ and Ensemble II with $s_3 > 0$. During

the simulation, the planar surface was forced to remain normal to the $\hat{s}_3$-axis, but allowed to move along this axis as in the first vibrational state of the hydrogen fluoride molecule. A simulation of 10,000 replicas over 100,000 time steps with $\Delta\tau = 0.2\,E_h^{-1}$ yielded an equilibrium value of 0.0334 ± 0.0003 for the displacement of the plane along the $\hat{s}_3$-axis and an energy of 1594.9 ± 2.0 cm^{-1}. This value is close to the converged variational value and indicates that the planar nodal surface is near the exact surface. However, some improvement in the approach is needed to get the correct value.

Since the local coordinates are not orthogonal, and the kinetic energy operator is not diagonal in the local coordinates, we expect the exact nodal surface to deviate from the planar approximation. The $|0,0,1\rangle$ simulation was repeated, incorporating the dynamic mode method based on a flexible nodal surface. The nodal surface was divided into a 32 x 32 rectangular grid. The array spanned the region $-1 < u < 2.5$ and $-2.5 < v < 2.5$, with displacements along s_3 associated with every grid point. A continuous surface was formed by linearly interpolating values for s_3 within the array panels. This characterization of the nodal surface retains some limitations as it does not allow for shapes such as folding. It is possible, however, to reparameterize the nodal surface during the simulation if necessary. Our parameterization does satisfy the necessary condition that there is one parameter, namely s_3, along which one ensemble is always expanded and the other always compressed. For this and all other symmetric states, the replica density was increased by recognising that the state is symmetric about $v = 0$ and running the simulation in the region $v > 0$. Any replica which moved to $v < 0$ was reflected about the plane $v = 0$.

The simulation proceeded, allowing 10,000 replicas to diffuse on either side of the nodal surface. After each replica was moved, its weight was multiplied first by the potential operator and secondly by the recross probability. It is time-consuming to calculate the recross probability for a replica near a complicated nodal surface. We made an approximate, yet

adequate, calculation of this probability in the following manner. The u and v coordinates of each replica locate it directly above or below, in the $\hat{s}_3$ direction, one panel of the nodal surface. Before a replica was moved, that panel was identified, and we determined on which side a replica was located. If the replica crossed the panel during the diffusion step, it was killed and its weight added to the flux crossing the panel in that direction. Otherwise, the distance of the replica to its associated panel was computed before and after the time step so that Eq. 8 could be applied, and the replica weight attenuated accordingly. Since the attenuated portion of the replica was lost due to its chance of crossing the nodal surface, the lost weight was also added to the estimated flux across the panel.

The replicas were allowed to diffuse for 100 time steps before the nodal surface was shifted. Over this time, the local energy of each ensemble was averaged, and the flux of replica weights from each side of every panel of the nodal surface was summed. As required by Eq. 22, an estimate of the population fraction was computed based upon the ratio of the total fluxes to either side of the node. The entire nodal surface was shifted by a uniform amount in s_3 based upon the difference in local energy across the barrier. Each panel was then shifted based upon the relative flux of replica weight on either side, leaving the average position of the node fixed in s_3. The simulation proceeded until the nodal surface reached equilibrium and the energy was calculated as usual. The final energy of the simulation was $1597.4 \pm 3.3\ cm^{-1}$ and the nodal surface is shown in Figure 3.

A point to note is that the flux of weights across a nodal panel in any direction can be estimated. When the panel is properly located, the nett flux through its surface is zero. In the simulation described above, the flux and recross corrections were calculated only if the contributing replica remained above or below the same panel over the course of its diffusion step. This approximation saved having to determine which panel a replica crossed, or which it was closest to, if it made a jump spanning one or more grid points. In practice, the grid points

were sufficiently separated that on average the replicas take many times steps to move from one panel to another.

The two excited states simulated so far fall within the class of solutions with a single, open, nodal surface. The $|0,0,1\rangle$ bend is the lowest energy state which can be simulated without any constraints other than the existence of an open nodal surface. The $|1,0,0\rangle_a$ antisymmetric stretch is the lowest energy state in its symmetry class which lies within this geometric class of nodal surfaces. It, too, could be simulated without approximation. We could simulate the $|1,0,0\rangle_s$ symmetric stretch as in the previous simulation, by starting two ensembles either side of a planar nodal surface constrained to be parallel to the plane $u = 0$. If this constraint on such a simulation is removed, the surface rotates into the uv plane and the wavefunction evolves to the lower energy $|0,0,1\rangle$ bend state described earlier. This result is not surprising as the constraints are then no different to those used to simulate the $|0,0,1\rangle$ state. The symmetric stretch state can be simulated by imposing the additional constraint that it must be orthogonal to the bend. However, for this simulation we return to the independent coordinate approximation and assume the nodal surface to be independent of s_3 and a function of u and v only.

The trial nodal surface was initially placed at $u = 0$. Figure 2 and symmetry arguments show the approximate surface to be some curve symmetric about $s_1 = s_2$. The nodal surface was defined by 40 points (u,v) equally spaced in the region $0 < v < 2$. Instead of being made up of rectangular panels, the panels were strips extending infinitely in s_3. The simulation was brought to equilibrium and the energy of the ensembles was then averaged over 30,000 iterations and found to be $3653.3 \pm 2.5\ cm^{-1}$. The positions of the panels were also averaged over the simulation and are shown in Figure 4. The solid line running through the surface points is the node of the pure Morse state defined by

$$|100\rangle_s = (|1\rangle|0\rangle|0\rangle + |0\rangle|1\rangle|0\rangle)\sqrt{2} = 0 \qquad (28)$$

in terms of the symmetry coordinates u and v. The benefit of the dynamic node technique can be seen in the deviation between the simulation and the pure Morse surfaces. The shift between the two is primarily due to the f_{12} coupling term; finding the position of the true nodal surface analytically is difficult, if not impossible.

IX. FUTURE DEVELOPMENTS

Before any quantum DMC simulation is started using the methods described in this paper, the basic geometry of the species of interest must be known. Such information includes the number of lobes in the wave function, and how many nodal surfaces divide these lobes topologically. A major limitation is that in general only the lowest energy states within each geometry can be calculated. This is not an issue for one dimensional problems, for which there can be only one wave function in each geometry, but in higher dimensions the situation becomes complicated. For example, in the full treatment of the water excitations, the class of wave functions with two lobes include the $|0,0,1\rangle$, $|1,0,0\rangle_s$, $|1,0,0\rangle_a$, $|2,0,0\rangle_s$ states, and possibly others. The $|0,0,1\rangle$ bend and the $|1,0,0\rangle_a$ antisymmetric stretch states are the lowest in energy of their respective symmetry classes and can be simulated directly. The $|1,0,0\rangle_s$ and $|2,0,0\rangle_s$ states cannot be simulated without appropriate orthogonality constraints with all other states of lower energy within that geometry.

A method for orthogonalisation was described by Coker and Watts [6]. They constructed histogram wave functions for lower states and computed the dot product between states by summing the product of histogram values over the wave functions. A major limitation of their approach is that the replica density must be high enough to significantly fill each element of the histogram. This is not a problem in low dimensions, but even with a modest 50 bins in

Table 1 Dynamic node Morse Oscillator Simulation Results

	Energy			Node	
State	Lobe	Average	Exact	DMC	Exact
$\|0\rangle$	$E_I = 0.4425 \pm 4.10^{-4}$	0.4425 ± 4.10^{-4}	0.44221	-	-
$\|1\rangle$	$E_I = 1.2968 \pm 5.10^{-4}$ $E_{II} = 1.2969 \pm 4.10^{-4}$	1.2968 ± 3.10^{-4}	1.29664	$N_1 = 0.2288 \pm 1.10^{-4}$	0.22876
$\|2\rangle$	$E_I = 2.1110 \pm 5.10^{-4}$ $E_{II} = 2.1111 \pm 7.10^{-4}$ $E_{III} = 2.1112 \pm 4.10^{-4}$	2.1111 ± 3.10^{-4}	2.11107	$N_1 = -0.3727 \pm 2.10^{-4}$ $N_2 = 1.1878 \pm 3.10^{-4}$	-0.37252 1.18821
$\|3\rangle$	$E_I = 2.8863 \pm 6.10^{-4}$ $E_{II} = 2.8864 \pm 9.10^{-4}$ $E_{III} = 2.8862 \pm 9.10^{-4}$ $E_{IV} = 2.8863 \pm 4.10^{-4}$	2.8863 ± 4.10^{-4}	2.8855	$N_1 = -0.7631 \pm 2.10^{-4}$ $N_2 = 0.5510 \pm 2.10^{-4}$ $N_3 = 1.9926 \pm 3.10^{-4}$	-0.76295 0.55133 1.99321

Table 2 HF dynamic-node DMC results

	Energy cm^{-1}			Node	
State $\|v,l\rangle$	Lobe	Average	Exact	DMC	Exact
$\|0,0\rangle$	$E_I = 2042.9 \pm 0.5$	2042.9 ± 0.5	2043.2	-	-
$\|0,1\rangle$	$E_I = 2085.4 \pm 1.1$ $E_{II} = 2085.8 \pm 1.3$	2085.6 ± 0.8	2085.1	$N_x = -0.001 \pm 0.006$	0.0000
$\|0,2\rangle$	$E_I = 2169.3 \pm 1.7$ $E_{II} = 2270.4 \pm 1.6$ $E_{III} = 2167.9 \pm 1.5$	2167.9 ± 0.9	2168.9	$N_x = -0.002 \pm 0.007$ $N_y = 0.003 \pm 0.006$	0.0000 0.0000
$\|1,0\rangle$	$E_I = 5970.5 \pm 2.1$ $E_{II} = 5970.0 \pm 1.7$	5970.3 ± 1.4	5970.2	$N_z = 1.77326 \pm 4.10^{-5}$	1.7732
$\|1,1\rangle$	$E_I = 6011.3 \pm 1.4$ $E_{II} = 6011.8 \pm 1.5$ $E_{III} = 6011.5 \pm 1.4$	6011.5 ± 0.8	6012.1	$N_x = -0.003 \pm 0.004$ $N_z = 1.77341 \pm 7.10^{-5}$	0.0000 1.7732

Table 3 Variational energies of the water monomer excited states, measured from the $|000\rangle$ ground state reference energy of 4623.02 cm^{-1}

State	Energy
$\|0,0,1\rangle$	1,595
$\|0,0,2\rangle$	3,152
$\|1,0,0\rangle_s$	3,656
$\|1,0,0\rangle_a$	3,755
$\|0,0,3\rangle$	4,663
$\|1,0,1\rangle_s$	5,252
$\|1,0,1\rangle_a$	5,343
$\|1,0,2\rangle_s$	6,807
$\|1,0,2\rangle_a$	6,889
$\|2,0,0\rangle_s$	7,201
$\|2,0,0\rangle_a$	7,250
$\|1,1,0\rangle$	7,445
$\|2,0,1\rangle_s$	8,762
$\|2,0,1\rangle_a$	8,807

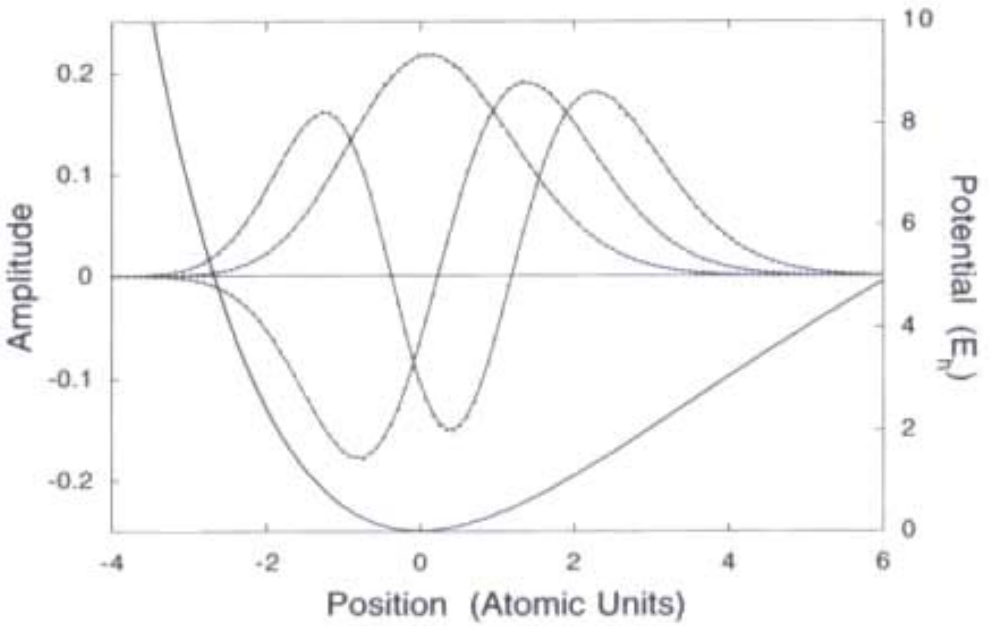

Figure 1 The first three Morse oscillator wave functions and the Morse potential

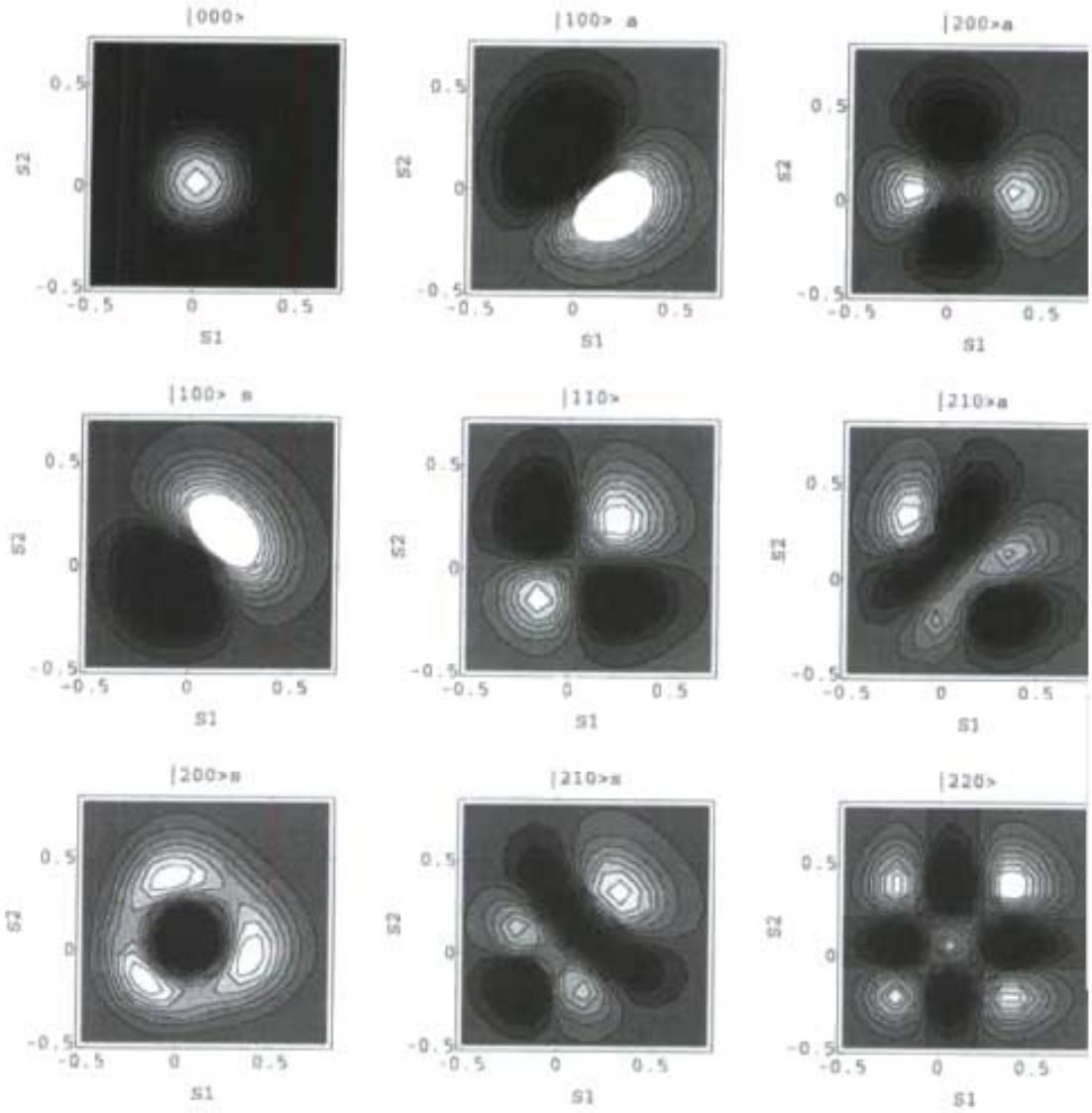

Figure 2 Approximate wave functions for the water molecule projected onto the s_1s_2 plane

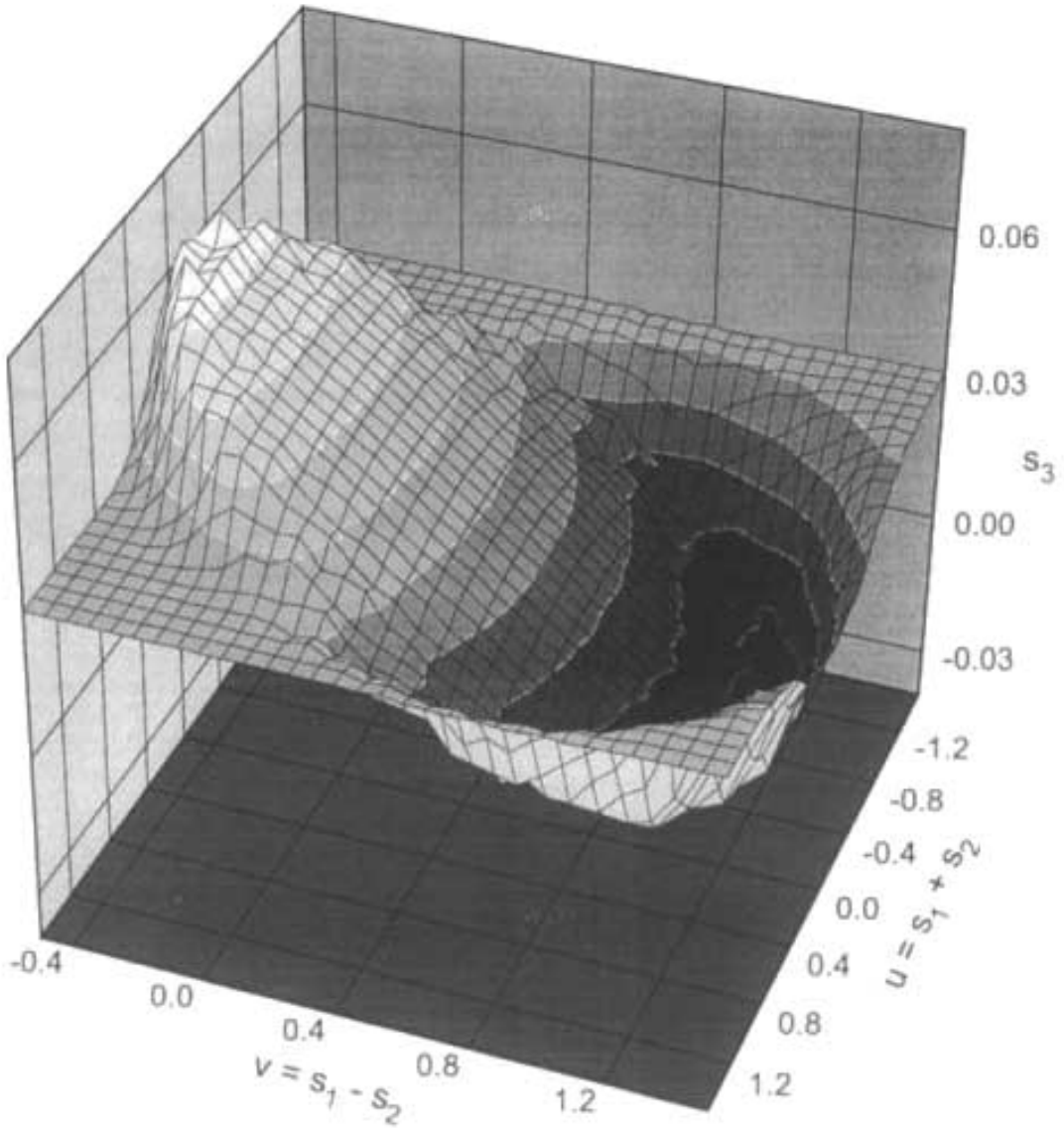

Figure 3 Nodal surface of the $|1,0,0\rangle$ bend state of the water molecule

each dimension, a four or five dimensional problem becomes impossible. This class of problems is precisely where DMC simulations excel over other methods of calculating quantum state energies. It seems clear that by combining orthogonality and geometrical constraints with flexible nodal surfaces, more progress can be made. There needs to be more work in this area to allow dynamic node DMC to simulate excited states of more complicated problems.

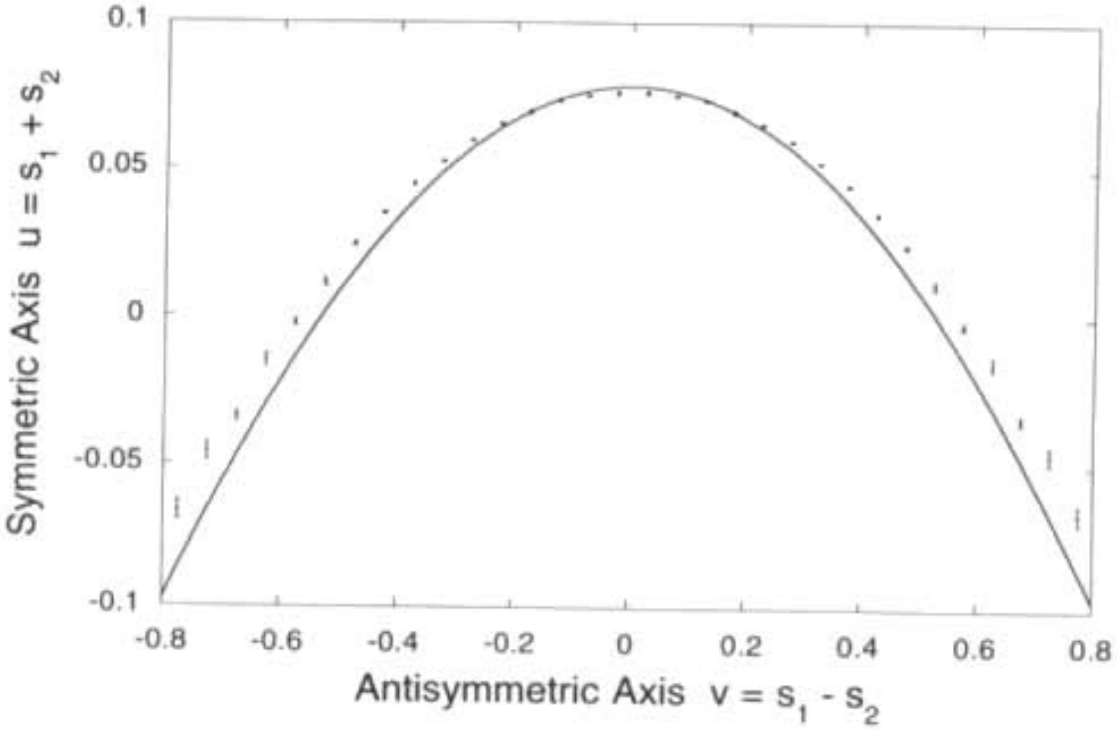

Figure 4 Nodal section of the $|1,0,0\rangle$ symmetric stretch of the water molecule showing how it is preturbed from the node of the pure Morse state

REFERENCES

[1] M. A. Suhm, and O. Watts, *Phys. Rep. 204,* 293 (1991).

[2] B. Anderson, J. *Chem. Phys. 63,* 1499 (1975); *65,* 4121 (1976); *73,* 3897 (1980); J. B. Anderson and B. H. Freihaut, J. *Comput. Phys. 31,* 425 (1979).

[3] M. Ceperley and B. Bernu, J. *Chem. Phys.89,* 6316 (1988); B. Bernu, D. M. Ceperley and W. A. Lester, J. *Chem. Phys. 93,* 552 (1990). M. Ceperley and B. J. Alder, J. *Chem. Phys. 81,* 5833 (1984).

[4] D. F. Coker and R. O. Watts, Mol. Phys. *58,* 1113 (1986).

[5] H. Sun and R. O. Watts, J. *Chem. Phys.* 92, 603 (1990); M. Quack and M. A. Suhm, Mol. Phys. *69,* 791 (1990).

[6] R. Reimers and R. O. Watts, *Mol. Phys.52,* 357 (1984).

[7] D. F. Coker, R. E. Miller, and R. O. Watts, J. *Chem. Phys. 82,* 3554 (1985).